



0590
JL5

RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/982,835
Source: OIPE
Date Processed by STIC: 6-25-02

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY
FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.
Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.
Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>>), EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. Hand Carry directly to:
U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7th Floor, Examiner Name, Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202
Or
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Does Not Comply
Corrected Diskette Needed



OIPE

RAW SEQUENCE LISTING DATE: 06/25/2002
PATENT APPLICATION: US/09/982,835 TIME: 10:06:43

Input Set : A:\g15047u2.txt
Output Set: N:\CRF3\06252002\I982835.raw

SEQUENCE LISTING

4 (1) GENERAL INFORMATION:
 6 (i) APPLICANT: RABIN, Mark B.
 8 (ii) TITLE OF INVENTION: MUTATIONS IN THE BRCA1 GENE
 10 (iii) NUMBER OF SEQUENCES: 10
 12 (iv) CORRESPONDENCE ADDRESS:
 13 (A) ADDRESSEE: Morgan, Lewis & Bockius LLP
 14 (B) STREET: 1111 Pennsylvania Avenue, N.W.
 15 (C) CITY: Washington
 16 (D) STATE: DC
 17 (E) COUNTRY: USA
 18 (F) ZIP: 20004
 20 (v) COMPUTER READABLE FORM:
 21 (A) MEDIUM TYPE: Diskette
 22 (B) COMPUTER: IBM Compatible
 23 (C) OPERATING SYSTEM: Windows
 24 (D) SOFTWARE: FastSEQ for Windows Version 2.0b
 26 (vi) CURRENT APPLICATION DATA:
 C--> 27 (A) APPLICATION NUMBER: US/09/982,835
 C--> 28 (B) FILING DATE: 17-Jun-2002
 34 (vii) PRIOR APPLICATION DATA:
 31 (A) APPLICATION NUMBER: US 09/038,946
 32 (B) FILING DATE: 1998-03-12
 35 (A) APPLICATION NUMBER: US 09/697,149
 36 (B) FILING DATE: 2000-10-27
 38 (viii) ATTORNEY/AGENT INFORMATION:
 39 (A) NAME: Michael S. Tuscan, Ph.D.
 40 (B) REGISTRATION NUMBER: 43,210
 41 (C) REFERENCE/DOCKET NUMBER: 44921-5047-02-US
 43 (ix) TELECOMMUNICATION INFORMATION:
 44 (A) TELEPHONE: 202-739-3000
 45 (B) TELEFAX: 202-739-3001

ERRORED SEQUENCES

48 (2) INFORMATION FOR SEQ ID NO: 1:
 50 (i) SEQUENCE CHARACTERISTICS:
 51 (A) LENGTH: 5710 base pairs
 52 (B) TYPE: nucleic acid
 53 (C) STRANDEDNESS: single
 54 (D) TOPOLOGY: linear
 57 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

Sequence is 5711
in length.

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/982,835

DATE: 06/25/2002
TIME: 10:06:43

Input Set : A:\gl5047u2.txt
Output Set: N:\CRF3\06252002\I982835.raw

59	AGCTCGCTGA	GACTTCCTGG	ACCCCGCACC	AGGCTGTGGG	GTTTCTCAGA	TAACTGGGCC	60
60	CCTCGCCTCA	GGAGGCCCTC	ACCCTCTGCT	CTGGGTAAAG	TTCATTGGAA	CAGAAAAGAAA	120
61	TGGATTATC	TGCTCTTCGC	GTTGAAGAAG	TACAAAATGT	CATTAATGCT	ATGCAGAAAA	180
62	TCTTAGAGTG	TCCCATCTGT	CTGGAGTTGA	TCAAGGAACC	TGTCTCCACA	AAGTGTGACC	240
63	ACATATTTC	CAAATTTGC	ATGCTGAAAC	TTCTCAACCA	GAAGAAAGGG	CCTTCACAGT	300
64	GTCCTTTATG	TAAGAATGAT	ATAACCAAAA	GGAGCCTACA	AGAAAGTAGC	AGATTTAGTC	360
65	AACTGTTGA	AGAGCTATTG	AAAATCATT	GTGCTTTCA	GCTTGACACA	GGTTTGGAGT	420
66	ATGCAAACAG	CTATAATTT	GCAAAAAAGG	AAAATAACTC	TCCTGAACAT	CTAAAAGATG	480
67	AAGTTCTAT	CATCCAAAGT	ATGGGCTACA	GAAACCGTGC	CAAAGACTT	CTACAGAGTG	540
68	AACCCGAAAA	TCCTTCCTG	CAGGAAACCA	GTCTCAGTGT	CCAACCTCT	AACCTTGGAA	600
69	CTGTGAGAAC	TCTGAGGACA	AAGCAGCGGA	TACAACCTCA	AAAGACGTCT	GTCTACATTG	660
70	AATTGGGATC	TGATTCTTCT	GAAGATAACG	TTAATAAGGC	AACTTATGTC	AGTGTGGGAG	720
71	ATCAAGAAC	GTTACAAATC	ACCCCTCAAG	GAACCAAGGG	TGAAATCAGT	TTGGATTCTG	780
72	CAAAAAAGGC	TGCTGTGAA	TTTCTGAGA	CGGATGTAAC	AAATACTGAA	CATCATCAAC	840
73	CCAGTAATAA	TGATTGAAAC	ACCACTGAGA	AGCAGTCAGC	TGAGAGGCAT	CCAGAAAAGT	900
74	ATCAGGGTAG	TTCTGTTCA	AACTTGATCG	TGGAGCCATG	TGGCACAAAT	ACTCATGCCA	960
75	GCTCATTACA	GCATGAGAAC	AGCAGTTAT	TACTCACTAA	AGACAGAATG	AATGTAGAAA	1020
76	AGGCTGAATT	CTGTAATAAA	ACCAAACAGC	CTGGCTTAGC	AAGGAGCCAA	CATAACAGAT	1080
77	GGGCTGGAAG	TAAGGAAACA	TGTAATGATA	GGCGGACTCC	CAGCACAGAA	AAAAAGGTAG	1140
78	ATCTGAATGC	TGATCCCCCTG	TGTGAGAGAA	AAGAATGGAA	TAAGCAGAAA	CTGCCATGCT	1200
79	CAAGAACATCC	TAGAGATACT	GAAGATGTT	CTTGATAAC	ACTAAATAGC	AGCATTCAAGA	1260
80	AAGTTAATGA	GTGGTTTCC	AGAAGTGTATG	AACTGTTAGG	TTCTGATGAC	TCACATGATG	1320
81	GGGAGTCTGA	ATCAAATGCC	AAAGTAGCTG	ATGTATTGGA	CGTTCTAAAT	GAGGTTAGATG	1380
82	AATTATTCTGG	TTCTTCAGAG	AAAATAGACT	TACTGGCCAG	TGATCCTCAT	GAGGCTTTAA	1440
83	TATGAAAAG	TGAAAGAGTT	CACTCCAAAT	CAGTAGAGAG	TAATATTGAA	GACAAAATAT	1500
84	TTGGGAAAAC	CTATCGGAAG	AAGGCAAGCC	TCCCCAACTT	AAGCCATGTA	ACTGAAAATC	1560
85	TAATTATAGG	AGCATTGTT	ACTGAGCCAC	AGATAATACA	AGAGCGTCCC	CTCACAAATA	1620
86	AATTAAAGCG	TAAAAGGAGA	CCTACATCAG	GCCTTCATCC	TGAGGATTT	ATCAAGAAAG	1680
87	CAGATTGGC	AGTCCAAAG	ACTCCTGAAA	TGATAAAATCA	GGGAACAAAC	CAAACGGAGC	1740
88	AGAATGGTCA	AGTGATGAAT	ATTACTAATA	GTGGTCATGA	GAATAAAACA	AAAGGTGATT	1800
89	CTATTCAAGAA	TGAGAAAAAT	CCTAACCCAA	TAGAATCACT	CGAAAAAGAA	TCTGCTTTCA	1860
90	AAACGAAAGC	TGAACCTATA	AGCAGCAGTA	TAAGCAATAT	GGAACTCGAA	TTAAATATCC	1920
91	ACAATTCAA	AGCACCTAAA	AAGAATAGGC	TGAGGAGGAA	GTCTTCTACC	AGGCATATTC	1980
92	ATGCGCTTGA	ACTAGTAGTC	AGTAGAAATC	TAAGCCCACC	TAATTGTACT	GAATTGCAAA	2040
93	TTGATAGTTG	TTCTAGCAGT	GAAGAGATAA	AGAAAAAAAG	GTACAACCAA	ATGCCAGTCA	2100
94	GGCACAGCAG	AAACCTACAA	CTCATGGAAG	GTAAAGAAC	TGCAACTGG	GCCAAGAAGA	2160
95	GTAAACAAAGCC	AAATGAACAG	ACAAGTAAAA	GACATGACAG	TGATACTTTC	CCAGAGCTGA	2220
96	AGTTAACAAA	TGCACCTGGT	TCTTTACTA	AGTGTCAAA	TACCACTGAA	CTTAAAGAAT	2280
97	TTGTCAATCC	TAGCCTTCCA	AGAGAAGAAA	AAGAAGAGAA	ACTAGAAACA	GTAAAGTGT	2340
98	CTAATAATGC	TGAAGACCCC	AAAGATCTCA	TGTTAACTGG	AGAAAGGGTT	TTGCAAACGT	2400
99	AAAGATCTGT	AGAGAGTAGC	AGTATTCAC	TGGTACCTGG	TACTGATTAT	GGCACTCAGG	2460
100	AAAGTATCTC	GTTACTGGAA	GTTAGCACTC	TAGGGAAGGC	AAAAACAGAA	CCAAATAAAAT	2520
101	GTGTGAGTCA	GTGTGCAAGCA	TTTGAACACC	CCAAGGGACT	AATTCTGGT	TGTTCCAAAG	2580
102	ATAATAGAAA	TGACACAGAA	GGCTTTAAGT	ATCCATTGGG	ACATGAAGTT	AACCACAGTC	2640
103	GGGAAACAAAG	CATAGAAATG	GAAGAAAGTG	AACTTGATGC	TCAGTATTG	CAGAATACAT	2700
104	TCAAGGTTTC	AAAGCGCCAG	TCATTTGCTC	TGTTTCAAA	TCCAGGAAT	GCAGAAGAGG	2760
105	AATGTGCAAC	ATTCTCTGCC	CACTCTGGGT	CCTTAAAGAA	ACAAAGTCCA	AAAGTCACCT	2820
106	TTGAATGTGA	ACAAAAGGAA	GAAAATCAAG	GAAAGAATGA	GTCTAATATC	AAGCCTGTAC	2880
107	AGACAGTAA	TATCACTGCA	GGCTTCCTG	TGGTTGGTCA	GAAAGATAAG	CCAGTTGATA	2940

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/982,835

DATE: 06/25/2002
TIME: 10:06:43

Input Set : A:\gl5047u2.txt
Output Set: N:\CRF3\06252002\I982835.raw

108	ATGCCAAATG	TAGTATCAA	GGAGGCTCTA	GGT	TTTGTCT	ATCATCTCAG	TTCAGAGGCA	3000
109	ACGAAACTGG	ACTCATTACT	CCAAATAAAC	ATGGACTTTT	ACAAAACCCA	TATCGTATAC	3060	
110	CACCACTTT	TCCCATCAAG	TCATTTGTTA	AAACTAAATG	TAAGAAAAAT	CTGCTAGAGG	3120	
111	AAAACCTTG	GAACATTCA	ATGTCACCTG	AAAGAGAAAT	GGGAAATGAG	AACATTCCAA	3180	
112	GTACAGTGAG	CACAAATTAGC	CGTAATAACA	TTAGAGAAAA	TGTTTTTAAA	GGAGCCAGCT	3240	
113	CAAGCAATAT	TAATGAAGTA	GGT	TCAGTCA	CTAATGAAGT	GGGCTCCAGT	3300	
114	TAGGTTCCAG	TGATGAAAC	ATTCAAGCAG	AACTAGGTAG	AAACAGAGGG	CCAAAATTGA	3360	
115	ATGCTATGCT	TAGATTAGGG	GTTTGCAC	CTGAGGTCTA	AAACAAAGT	CTTCCTGGAA	3420	
116	GTAATTGAA	GCATCCTGAA	ATAAAAAAAGC	AAAGATATGA	AGAAGTAGT	CAGACTGTTA	3480	
117	ATACAGATT	CTCTCCATAT	CTGATTTCA	ATAACTT	AGAC	CCATATG	3540	
118	ATGCATCTCA	GGTTGTTCT	GAGACACCTG	ATGACCTGTT	AGATGATGGT	GAAATAAAGG	3600	
119	AAGATACTAG	TTTGCTGAA	AATGACATTA	AGGAAAGTTC	TGCTGTTTT	AGCAAAAGCG	3660	
120	TCCAGAGGAG	AGAGCTTAGC	AGGAGTCTA	GCCCTTTCAC	CCATACACAT	TTGGCTCAGG	3720	
121	GTTACCGAAG	AGGGGCCAAG	AAATTAGAGT	CCTCAGAAGA	GAAC	TTATCT	AGTGAGGATG	3780
122	AAGAGCTCC	CTGCTTCAA	CACTTGTAT	TTGGTAAAGT	AAACAATATA	CCTTCTCAGT	3840	
123	CTACTAGGCA	TAGCACCGTT	GCTACCGAGT	GTCTGTCTAA	GAACACAGAG	GAGAATTAT	3900	
124	TATCATTGAA	GAATAGCTTA	AATGACTGCA	GTAACCAGGT	AATATTGCA	AAGGCATCTC	3960	
125	AGGAACATCA	CCTTAGTGAG	GAAACAAAAT	GTTCTGCTAG	CTTGT	TTCT	TCACAGTGCA	4020
126	GTGAATTGGA	AGACTTGACT	GCAAATACAA	ACACCCAGGA	TCC	TTCTTG	ATTGGTTCTT	4080
127	CCAAACAAAT	GAGGCATCAG	TCTGAAAGCC	AGGGAGTTGG	TCTGAGTGAC	AAGGAATTGG	4140	
128	TTTCAGATGA	TGAAGAAAGA	GGAACGGGCT	TGGAAGAAAA	TAATCAAGAA	GAGCAAAGCA	4200	
129	TGGATTCAA	CTTAGGTGAA	GCAGC	ATCTG	GGTGTGAGAG	TGAAACAAAGC	4260	
130	ACTGCTCAGG	GCTATCCTCT	CAGAGTGACA	TTTTAAC	AC	CC	TCAGCAGAGG	4320
131	AACATAACCT	GATAAAGCTC	CAGCAGGAAA	TGGCTGA	AACT	GTG	TTAGAACAGC	4380
132	ATGGGAGCCA	GCCTTCTAAC	AGCTACCC	TT	CCATCATAAG	TGACTCCTCT	GCCCTTGAGG	4440
133	ACCTGCGAAA	TCCAGAACAA	AGCACATCAG	AAAAAGCAGT	ATTAAC	TTCA	CAGAAAAGTA	4500
134	GTGAATACCC	TATAAGCCAG	AATCCAGAAG	GCCTT	CTGC	TGACAAG	TT	4560
135	CAGATAGTTC	TACCA	GTAA	AGC	TTGGA	AAGGT	CATCC	4620
136	GCCC	CATT	AGATGATAGG	TGGTACATGC	ACAGTTGCTC	TGGGAG	CTT	4680
137	ACTACCCATC	TCAAGAGGAG	CTCATT	AAAG	TTGTTGATGT	GGAGGAGC	AA	4740
138	AGTCTGGGCC	ACACGATTG	ACGGAAACAT	CTTACTTGCC	AAGGCAAGAT	CTAGAGGGAA	4800	
139	CCCCTTACCT	GGAA	ATCTGGA	ATCAGCCTCT	TCTCTGATGA	CCCTGA	ATCT	4860
140	AAGACAGAGC	CCCAGAGTC	GCTCGTGTG	GCAACATACC	ATCTTCAACC	TCTGCATTGA	4920	
141	AAGTTCCCCA	ATTGAAAGT	GCAGAATCTG	CCCAGGGTCC	AGCTGCTGCT	CATACTACTG	4980	
142	ATACTGCTGG	GTATAATGCA	ATGGAAGAAA	GTGTGAGCAG	GGAGAAGCCA	GAATTGACAG	5040	
143	CTTCAACAGA	AAGGGTCAAC	AAAAGAATGT	CCATGGTGGT	GTCTGGCCTG	ACCCAGAAG	5100	
144	AATT	TATGCT	CGTGTACAAG	TTTGCAGAA	AACACCACAT	CACTTTA	ACTAATT	5160
145	CTGAAGAGAC	TACTCATGTT	GT	TATGAA	CAGATGCTGA	GT	TTGTGT	5220
146	TGAAATATT	TCTAGGAATT	GC	GGGAGGAA	AA	TGGGTAGT	TAGCTATT	5280
147	AGTCTATTA	AGAAAGAAA	ATGCTGA	ATGATGATT	TGAAGTCAGA	GGAGATGT	5340	
148	TCAATGGAAG	AAACCACCAA	GGT	CCAAAGC	GAGCAAGAGA	ATCCCAGGAC	AGAAAGATCT	5400
149	TCAGGGGCT	AGAAATCTGT	TGCTATGGC	CCTCACC	CATGCC	CACA	GATCAACTG	5460
150	AATGGGATGGT	ACAGCTGTGT	GGTGCTTCTG	TGGTGAAGGA	GCTT	TCATCA	TTCA	5520
151	GCACAGGTGT	CCACCCAA	TG	GGTTGTG	AGCCAGATGC	CTGGACAGAG	GACAATGGCT	5580
152	TCCATGCAAT	TGGCAGATG	TGTGAGGCAC	CTGTG	GTGAC	CCGAGAGT	GG	5640
153	GTGTAGCACT	CTACAGTGC	CAGGAGCTGG	ACAC	CTAC	GTGAC	ATCCCCCACA	5700
154	GCCACTACTG	A						5710

(2) INFORMATION FOR SEQ ID NO: 2:
(i) SEQUENCE CHARACTERISTICS:

Sequence is 5710

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/982,835

DATE: 06/25/2002
TIME: 10:06:43

Input Set : A:\gl5047u2.txt
Output Set: N:\CRF3\06252002\I982835.raw

159 (A) LENGTH: 1863 amino acids
160 (B) TYPE: amino acid
161 (C) STRANDEDNESS: single
162 (D) TOPOLOGY: linear
164 (ii) MOLECULE TYPE: protein
166 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
168 Met Asp Leu Ser Ala Leu Arg Val Glu Val Gln Asn Val Ile Asn
169 1 5 10 15
170 Ala Met Gln Lys Ile Leu Glu Cys Pro Ile Cys Leu Glu Leu Ile Lys
171 20 25 30
172 Glu Pro Val Ser Thr Lys Cys Asp His Ile Phe Cys Lys Phe Cys Met
173 35 40 45
174 Leu Lys Leu Leu Asn Gln Lys Lys Gly Pro Ser Gln Cys Pro Leu Cys
175 50 55 60
176 Lys Asn Asp Ile Thr Lys Arg Ser Leu Gln Glu Ser Thr Arg Phe Ser
177 65 70 75 80
178 Gln Leu Val Glu Glu Leu Leu Lys Ile Ile Cys Ala Phe Gln Leu Asp
179 85 90 95
180 Thr Gly Leu Glu Tyr Ala Asn Ser Tyr Asn Phe Ala Lys Lys Glu Asn
181 100 105 110
182 Asn Ser Pro Glu His Leu Lys Asp Glu Val Ser Ile Ile Gln Ser Met
183 115 120 125
184 Gly Tyr Arg Asn Arg Ala Lys Arg Leu Leu Gln Ser Glu Pro Glu Asn
185 130 135 140
186 Pro Ser Leu Gln Glu Thr Ser Leu Ser Val Gln Leu Ser Asn Leu Gly
187 145 150 155 160
188 Thr Val Arg Thr Leu Arg Thr Lys Gln Arg Ile Gln Pro Gln Lys Thr
189 165 170 175
190 Ser Val Tyr Ile Glu Leu Gly Ser Asp Ser Ser Glu Asp Thr Val Asn
191 180 185 190
192 Lys Ala Thr Tyr Cys Ser Val Gly Asp Gln Glu Leu Leu Gln Ile Thr
193 195 200 205
194 Pro Gln Gly Thr Arg Asp Glu Ile Ser Leu Asp Ser Ala Lys Lys Ala
195 210 215 220
196 Ala Cys Glu Phe Ser Glu Thr Asp Val Thr Asn Thr Glu His His Gln
197 225 230 235 240
198 Pro Ser Asn Asn Asp Leu Asn Thr Thr Glu Lys Arg Ala Ala Glu Arg
199 245 250 255
200 His Pro Glu Lys Tyr Gln Gly Ser Ser Val Ser Asn Leu His Val Glu
201 260 265 270
202 Pro Cys Gly Thr Asn Thr His Ala Ser Ser Leu Gln His Glu Asn Ser
203 275 280 285
204 Ser Leu Leu Leu Thr Lys Asp Arg Met Asn Val Glu Lys Ala Glu Phe
205 290 295 300
206 Cys Asn Lys Ser Lys Gln Pro Gly Leu Ala Arg Ser Gln His Asn Arg
207 305 310 315 320
208 Trp Ala Gly Ser Lys Glu Thr Cys Asn Asp Arg Arg Arg Thr Pro Ser Thr
209 325 330 335
210 Glu Lys Lys Val Asp Leu Asn Ala Asp Pro Leu Cys Glu Arg Lys Glu

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/982,835

DATE: 06/25/2002
TIME: 10:06:43

Input Set : A:\gl5047u2.txt
Output Set: N:\CRF3\06252002\I982835.raw

211	340	345	350
212	Trp Asn Lys Gln Lys Leu Pro Cys Ser Glu Asn Pro Arg Asp Thr Glu		
213	355	360	365
214	Asp Val Pro Trp Ile Thr Leu Asn Ser Ser Ile Gln Lys Val Asn Glu		
215	370	375	380
216	Trp Phe Ser Arg Ser Asp Glu Leu Leu Gly Ser Asp Asp Ser His Asp		
217	385	390	395
218	Gly Glu Ser Glu Ser Asn Ala Lys Val Ala Asp Val Leu Asp Val Leu		
219	405	410	415
220	Asn Glu Val Asp Glu Tyr Ser Gly Ser Ser Glu Lys Ile Asp Leu Leu		
221	420	425	430
222	Ala Ser Asp Pro His Glu Ala Leu Ile Cys Lys Ser Glu Arg Val His		
223	435	440	445
224	Ser Lys Ser Val Glu Ser Asn Ile Glu Asp Lys Ile Phe Gly Lys Thr		
225	450	455	460
226	Tyr Arg Lys Lys Ala Ser Leu Pro Asn Leu Ser His Val Thr Glu Asn		
227	465	470	475
228	Leu Ile Ile Gly Ala Phe Val Thr Glu Pro Gln Ile Ile Gln Glu Arg		
229	485	490	495
230	Pro Leu Thr Asn Lys Leu Lys Arg Lys Arg Arg Pro Thr Ser Gly Leu		
231	500	505	510
232	His Pro Glu Asp Phe Ile Lys Lys Ala Asp Leu Ala Val Gln Lys Thr		
233	515	520	525
234	Pro Glu Met Ile Asn Gln Gly Thr Asn Gln Thr Glu Gln Asn Gly Gln		
235	530	535	540
236	Val Met Asn Ile Thr Asn Ser Gly His Glu Asn Lys Thr Lys Gly Asp		
237	545	550	555
238	Ser Ile Gln Asn Glu Lys Asn Pro Asn Pro Ile Glu Ser Leu Glu Lys		
239	565	570	575
240	Glu Ser Ala Phe Lys Thr Lys Ala Glu Pro Ile Ser Ser Ser Ile Ser		
241	580	585	590
242	Asn Met Glu Leu Glu Leu Asn Ile His Asn Ser Lys Ala Pro Lys Lys		
243	595	600	605
244	Asn Arg Leu Arg Arg Lys Ser Ser Thr Arg His Ile His Ala Leu Glu		
245	610	615	620
246	Leu Val Val Ser Arg Asn Leu Ser Pro Pro Asn Cys Thr Glu Leu Gln		
247	625	630	635
248	Ile Asp Ser Cys Ser Ser Ser Glu Glu Ile Lys Lys Lys Lys Tyr Asn		
249	645	650	655
250	Gln Met Pro Val Arg His Ser Arg Asn Leu Gln Leu Met Glu Gly Lys		
251	660	665	670
252	Glu Pro Ala Thr Gly Ala Lys Lys Ser Asn Lys Pro Asn Glu Gln Thr		
253	675	680	685
254	Ser Lys Arg His Asp Ser Asp Thr Phe Pro Glu Leu Lys Leu Thr Asn		
255	690	695	700
256	Ala Pro Gly Ser Phe Thr Lys Cys Ser Asn Thr Ser Glu Leu Lys Glu		
257	705	710	715
258	Phe Val Asn Pro Ser Leu Pro Arg Glu Glu Lys Glu Glu Lys Leu Glu		
259	725	730	735

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/982,835

DATE: 06/25/2002
TIME: 10:06:43

Input Set : A:\g15047u2.txt
Output Set: N:\CRF3\06252002\I982835.raw

260 Thr Val Lys Val Ser Asn Asn Ala Glu Asp Pro Lys Asp Leu Met Leu
 261 740 745 750
 262 Ser Gly Glu Arg Val Leu Gln Thr Glu Arg Ser Val Glu Ser Ser Ser
 263 755 760 765
 264 Ile Ser Leu Val Pro Gly Thr Asp Tyr Gly Thr Gln Glu Ser Ile Ser
 265 770 775 780
 266 Leu Leu Glu Val Ser Thr Leu Gly Lys Ala Lys Thr Glu Pro Asn Lys
 267 785 790 795 800
 268 Cys Val Ser Gln Cys Ala Ala Phe Glu Asn Pro Lys Gly Leu Ile His
 269 805 810 815
 270 Gly Cys Ser Lys Asp Asn Arg Asn Asp Thr Glu Gly Phe Lys Tyr Pro
 271 820 825 830
 272 Leu Gly His Glu Val Asn His Ser Arg Glu Thr Ser Ile Glu Met Glu
 273 835 840 845
 274 Glu Ser Glu Leu Asp Ala Gln Tyr Leu Gln Asn Thr Phe Lys Val Ser
 275 850 855 860
 276 Lys Arg Gln Ser Phe Ala Leu Phe Ser Asn Pro Gly Asn Ala Glu Glu
 277 865 870 875 880
 278 Glu Cys Ala Thr Phe Ser Ala His Ser Gly Ser Leu Lys Lys Gln Ser
 279 885 890 895
 280 Pro Lys Val Thr Phe Glu Cys Glu Gln Lys Glu Glu Asn Gln Gly Lys
 281 900 905 910
 282 Asn Glu Ser Asn Ile Lys Pro Val Gln Thr Val Asn Ile Thr Ala Gly
 283 915 920 925
 284 Phe Pro Val Val Gly Gln Lys Asp Lys Pro Val Asp Asn Ala Lys Cys
 285 930 935 940
 286 Ser Ile Lys Gly Gly Ser Arg Phe Cys Leu Ser Ser Gln Phe Arg Gly
 287 945 950 955 960
 288 Asn Glu Thr Gly Leu Ile Thr Pro Asn Lys His Gly Leu Leu Gln Asn
 289 965 970 975
 290 Pro Tyr Arg Ile Pro Pro Leu Phe Pro Ile Lys Ser Phe Val Lys Thr
 291 980 985 990
 292 Lys Cys Lys Lys Asn Leu Leu Glu Glu Asn Phe Glu Glu His Ser Met
 293 995 1000 1005
 294 Ser Pro Glu Arg Glu Met Gly Asn Glu Asn Ile Pro Ser Thr Val Ser
 295 1010 1015 1020
 296 Thr Ile Ser Arg Asn Asn Ile Arg Glu Asn Val Phe Lys Gly Ala Ser
 E--> 297 1025 1030 1035 104
 298 Ser Ser Asn Ile Asn Glu Val Gly Ser Ser Thr Asn Glu Val Gly Ser
 299 1045 1050 1055
 300 Ser Ile Asn Glu Ile Gly Ser Ser Asp Glu Asn Ile Gln Ala Glu Leu
 301 1060 1065 1070
 302 Gly Arg Asn Arg Gly Pro Lys Leu Asn Ala Met Leu Arg Leu Gly Val
 303 1075 1080 1085
 304 Leu Gln Pro Glu Val Tyr Lys Gln Ser Leu Pro Gly Ser Asn Cys Lys
 305 1090 1095 1100
 306 His Pro Glu Ile Lys Lys Gln Glu Tyr Glu Glu Val Val Gln Thr Val
 E--> 307 1105 1110 1115 112
 308 Asn Thr Asp Phe Ser Pro Tyr Leu Ile Ser Asp Asn Leu Glu Gln Pro

move number
 I space to left or
 use less amino
 acids per line

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/982,835

DATE: 06/25/2002
TIME: 10:06:43

Input Set : A:\gl5047u2.txt
Output Set: N:\CRF3\06252002\I982835.raw

309	1125	1130	1135
310	Met Gly Ser Ser His Ala Ser Gln Val Cys Ser Glu Thr Pro Asp Asp		
311	1140	1145	1150
312	Leu Leu Asp Asp Gly Glu Ile Lys Glu Asp Thr Ser Phe Ala Glu Asn		
313	1155	1160	1165
314	Asp Ile Lys Glu Ser Ser Ala Val Phe Ser Lys Ser Val Gln Arg Gly		
315	1170	1175	1180
316	Glu Leu Ser Arg Ser Pro Ser Pro Phe Thr His Thr His Leu Ala Gln		
E--> 317	1185	1190	1195
			120
318	Gly Tyr Arg Arg Gly Ala Lys Lys Leu Glu Ser Ser Glu Glu Asn Leu		
319	1205	1210	1215
320	Ser Ser Glu Asp Glu Glu Leu Pro Cys Phe Gln His Leu Leu Phe Gly		
321	1220	1225	1230
322	Lys Val Asn Asn Ile Pro Ser Gln Ser Thr Arg His Ser Thr Val Ala		
323	1235	1240	1245
324	Thr Glu Cys Leu Ser Lys Asn Thr Glu Glu Asn Leu Leu Ser Leu Lys		
325	1250	1255	1260
326	Asn Ser Leu Asn Asp Cys Ser Asn Gln Val Ile Leu Ala Lys Ala Ser		
E--> 327	1265	1270	1275
			128
328	Gln Glu His His Leu Ser Glu Glu Thr Lys Cys Ser Ala Ser Leu Phe		
329	1285	1290	1295
330	Ser Ser Gln Cys Ser Glu Leu Glu Asp Leu Thr Ala Asn Thr Asn Thr		
331	1300	1305	1310
332	Gln Asp Pro Phe Leu Ile Gly Ser Ser Lys Gln Met Arg His Gln Ser		
333	1315	1320	1325
334	Glu Ser Gln Gly Val Gly Leu Ser Asp Lys Glu Leu Val Ser Asp Asp		
335	1330	1335	1340
336	Glu Glu Arg Gly Thr Gly Leu Glu Glu Asn Asn Gln Glu Glu Gln Ser		
E--> 337	1345	1350	1355
			136
338	Met Asp Ser Asn Leu Gly Glu Ala Ala Ser Gly Cys Glu Ser Glu Thr		
339	1365	1370	1375
340	Ser Val Ser Glu Asp Cys Ser Gly Leu Ser Ser Gln Ser Asp Ile Leu		
341	1380	1385	1390
342	Thr Thr Gln Gln Arg Asp Thr Met Gln His Asn Leu Ile Lys Leu Gln		
343	1395	1400	1405
344	Gln Glu Met Ala Glu Leu Glu Ala Val Leu Glu Gln His Gly Ser Gln		
345	1410	1415	1420
346	Pro Ser Asn Ser Tyr Pro Ser Ile Ile Ser Asp Ser Ser Ala Leu Glu		
E--> 347	1425	1430	1435
			144
348	Asp Leu Arg Asn Pro Glu Gln Ser Thr Ser Glu Lys Ala Val Leu Thr		
349	1445	1450	1455
350	Ser Gln Lys Ser Ser Glu Tyr Pro Ile Ser Gln Asn Pro Glu Gly Leu		
351	1460	1465	1470
352	Ser Ala Asp Lys Phe Glu Val Ser Ala Asp Ser Ser Thr Ser Lys Asn		
353	1475	1480	1485
354	Lys Glu Pro Gly Val Glu Arg Ser Ser Pro Ser Lys Cys Pro Ser Leu		
355	1490	1495	1500
356	Asp Asp Arg Trp Tyr Met His Ser Cys Ser Gly Ser Leu Gln Asn Arg		
E--> 357	1505	1510	1515
			152

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/982,835

DATE: 06/25/2002
TIME: 10:06:43

Input Set : A:\gl5047u2.txt
Output Set: N:\CRF3\06252002\I982835.raw

358 Asn Tyr Pro Ser Gln Glu Glu Leu Ile Lys Val Val Asp Val Glu Glu
359 1525 1530 1535
360 Gln Gln Leu Glu Glu Ser Gly Pro His Asp Leu Thr Glu Thr Ser Tyr
361 1540 1545 1550
362 Leu Pro Arg Gln Asp Leu Glu Gly Thr Pro Tyr Leu Glu Ser Gly Ile
363 1555 1560 1565
364 Ser Leu Phe Ser Asp Asp Pro Glu Ser Asp Pro Ser Glu Asp Arg Ala
365 1570 1575 1580
366 Pro Glu Ser Ala Arg Val Gly Asn Ile Pro Ser Ser Thr Ser Ala Leu
E--> 367 1585 1590 1595 160
368 Lys Val Pro Gln Leu Lys Val Ala Glu Ser Ala Gln Gly Pro Ala Ala
369 1605 1610 1615
370 Ala His Thr Thr Asp Thr Ala Gly Tyr Asn Ala Met Glu Glu Ser Val
371 1620 1625 1630
372 Ser Arg Glu Lys Pro Glu Leu Thr Ala Ser Thr Glu Arg Val Asn Lys
373 1635 1640 1645
374 Arg Met Ser Met Val Val Ser Gly Leu Thr Pro Glu Glu Phe Met Leu
375 1650 1655 1660
376 Val Tyr Lys Phe Ala Arg Lys His His Ile Thr Leu Thr Asn Leu Ile
E--> 377 1665 1670 1675 168
378 Thr Glu Glu Thr Thr His Val Val Met Lys Thr Asp Ala Glu Phe Val
379 1685 1690 1695
380 Cys Glu Arg Thr Leu Lys Tyr Phe Leu Gly Ile Ala Gly Gly Lys Trp
381 1700 1705 1710
382 Val Val Ser Tyr Phe Trp Val Thr Gln Ser Ile Lys Glu Arg Lys Met
383 1715 1720 1725
384 Leu Asn Glu His Asp Phe Glu Val Arg Gly Asp Val Val Asn Gly Arg
385 1730 1735 1740
386 Asn His Gln Gly Pro Lys Arg Ala Arg Glu Ser Gln Asp Arg Lys Ile
E--> 387 1745 1750 1755 176
388 Phe Arg Gly Leu Glu Ile Cys Cys Tyr Gly Pro Phe Thr Asn Met Pro
389 1765 1770 1775
390 Thr Asp Gln Leu Glu Trp Met Val Gln Leu Cys Gly Ala Ser Val Val
391 1780 1785 1790
392 Lys Glu Leu Ser Ser Phe Thr Leu Gly Thr Gly Val His Pro Ile Val
393 1795 1800 1805
394 Val Val Gln Pro Asp Ala Trp Thr Glu Asp Asn Gly Phe His Ala Ile
395 1810 1815 1820
396 Gly Gln Met Cys Glu Ala Pro Val Val Thr Arg Glu Trp Val Leu Asp
E--> 397 1825 1830 1835 184
398 Ser Val Ala Leu Tyr Gln Cys Gln Glu Leu Asp Thr Tyr Leu Ile Pro
399 1845 1850 1855
400 Gln Ile Pro His Ser His Tyr
401 1860

See page 6

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/982,835

DATE: 06/25/2002
TIME: 10:06:44

Input Set : A:\g15047u2.txt
Output Set: N:\CRF3\06252002\I982835.raw

L:27 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:28 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:154 M:254 E: No. of Bases conflict, Input:5710 Counted:5711 SEQ:1
L:154 M:204 E: No. of Bases differ, LENGTH:Input:5710 Counted:5711 SEQ:1
L:297 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:2
M:332 Repeated in SeqNo=2